DETECTION OF STRUCTURAL ABNORMALITIES USING NEURAL NETS

M. Zak, A. Maccalla, V. Daggumati, S. Gulati, and N. Toomarian

Jet Propulsion Laboratory, Caltech

Abstract

This paper describes a feed forward neural net approach for detection of abnormal system behavior based upon sensor data analyses. A new dynamical invariant representing structural parameters of the system is introduced in such a way that any structural abnormalities in the system behavior are detected from the corresponding changes to the invariant Potential for application of this approach is discussed by analysis of results from telemetry monitoring of spacecraft systems for NASA interplanetary operations.

1. Introduction

Sensor data is a key source of information concerning the state and operational performance of the underlying system. When this data is presented in the form of a time series, two important problems can be posed: a) predicting future values of a time series from current and past values, and b) detection of abnormal behavior in the underlying system. In most of the existing methodologies[1,2], a solution to the second problem is based upon the solution to the first one: abnormal behavior is detected by analysis of the difference between the recorded and the expected/predicted values of the time series. Usually such an analysis is based upon comparing certain patterns such as the average value, average slope, average noise level, the period and phase of oscillations, and the frequency spectrum. Although all these characteristics may have some physical meaning when a time series represents a certain message in signal processing, they can change sharply in time when a time series describes the evolution of an underlying dynamical system such as the power system of a spacecraft or the catalytic converter of a car. Indeed, in the latter cases, a time series does not transmit any "man-made" message, and therefore, it may trove different dynamical invariant, In other words, it is reasonable to assume that when a lime series is describing the evolution of a dynamical system, its invariant can be represented by the coefficients of the differential (or the time-delay) equation which simulates the' dynamical process. Based upon this idea, we have developed the following strategy for detection of structural abnormalities: a) build a dynamical model which simulates a given time series, b) develop dynamical invariants whose change manifests structural abnormalities.

2. Dynamical Model

In this paper we deal with the sensor data in the form of a time series which describes the evolution of an underlying dynamical system. It will be assumed that this time series can not be approximated by a simple analytical expression, and therefore, it can be considered as a realization of an underlying stochastic process which can be described only in terms of a probability distribution. However, any information about this distribution can not be obtained from a simple realization of a stochastic process unless this process is stationary. (Then averaging over ensemble can be replaced by averaging over time))., But an assumption about the stationarity of the underlying stochastic process would exclude from consideration such important components of the dynamical process as linear and polinomial, trends, or harmonic oscillations. That is why we have to deal with non-stationary processes.

Our approach to building;; dynamical model is based upon progress in three independent fields: nonlinear dynamics, theory of stochastic processes, and artificial neural networks.

From the field of nonlinear dynamics, based upon the Takens theorem[3], any dynamical system which converges to an attractor of a lower (than original) dimensionality can be simulated (with a prescribed accuracy) by a time-delay equation

$$x(t) = F[x(t-\tau), x(t-2\tau), \ldots, x(t-m\tau)] \tag{1},$$

in which x(t) represents a given time series, and τ = constant is the time delay.

It was proven that the solution to Eq.(1) subject to appropriate initial conditions converges to the original time series:

$$x(t) = x(t_1), x(t_2)....etc \qquad (2)$$

when m in(I) is sufficiently large.

However, the function f', as well as the constant τ and m, are not specified by this theorem.

But the most "damaging" limitation of the model (1) is that the original time series must be stationary, since it represents an attractor. 1 his impies that for non-stationary time series the solution to (1) r-nay not converge to (2) at all.

Previously statistitions have developed a different approach to the same problem[4]: they approximated a stochastic process by a linear autoregressive model:

$$x(t) = \alpha_1 x(t - 1) + \alpha_2 x(t \quad ?)-:. \quad + \alpha_r (x_t - n) + N \qquad (3)$$

where a_i are constants, and N represents the contribution of a noise.

On first sight, Eq.(3) appears as a particular case of Eq.(1) when F is replaced by a linear function, and τ = 1. However, it actually has an important advantage over Eq.(1): it does not enforce stationarity of the time series (2). To be more precise, it requires certain transformations of (2) before the model can be applied. These transformations are supposed to "stationariz.e" the original time series. These types of transformations follow from the fact that the conditions of stationarity of the solution to Eq (3) coincide with the conditions of its stability, i.e. the process is non-stationary when

$$|G_i| > 1 \tag{4}$$

where G_i are the roots of the characteristic equation associated with Eq.(3).

The case $|G_i| > 1$ is usually excluded from considerations since it corresponds to an exponential instability which is unrealistic in physical systems under observation. However, the case $|G_i| = 1$ is realistic. Real and complex conjugates G_i incorporate trend and seasonal components, respectively, into the time series (2).

By applying a difference operator:

$$\nabla x_t = x_t - x_{t-1} = (1 - B)x_t \tag{5}$$

several times (where B is the shift operator), one can eliminate the trend from the time series:

$$x_t,\ x_{t-1}, \ldots etc \tag{6}$$

By applying a seasonal difference operator.

$$\nabla_s x_t = (1 - B^s)x_t = x_{t-s} \tag{7}$$

one can eliminate the seasonal components from the time series(6).

Unfortunately, it is not known in advance how many times the operators (5) or (7) should be applied to the original time series (6) for their stationarization. Moreover, in (7) the period S of the seasonal difference operator is also not prescribed. In the next section we will discuss possible ways to deal with these problems. Assuming that the time series (6) is stationarized, one can apply to them the model (1):

$$y(t) = F[y(t-1),\ y(t-2), \ldots, y(t-m)] \tag{8}$$

where

$$y_t, y_{t-1}, \ldots etc;\ (y_t = x_t - x_{t-\tau}) \tag{9}$$

are transformed series (6), and $\tau = 1$,

After fitting the model (8) to the time series (6), one can return to the old variable x(t) by exploiting the inverse operators $(1-B)^{-1}$ and $(1-B^s)^{1}$. For instance, if the stationarization procedure is performed by the operator (5), then:

$$x(t) = x(t-1) + F\{[x(t-1) - x(t-2)],\ [x(t-2) - x(t-3)] \ldots etc\} \tag{10}$$

Eq.(10) can be utilized for predictions of future values of (6), as well as for detections of structural abnormalities. However, despite the fact that Eq.(8) and (10) may be significantly different, their structure is uniquely defined by the same function F. Therefore, structural abnormalities which cause changes of the function F, can also be detected from Eq.(8) and consequently for that particular purpose the transition to Eq. (10) is not necessary.

It should be noticed that, strictly speaking the application of the stationarization procedure (5) and (7) to the time series (6) are justified only if the underlying model is linear. However, stationarity criteria nonlinear equations are,. more complex than for linear ones, in the same way as the criteria of stability are. Nevertheless, there are numerical evidences that even in nonlinear cases, the procedures (5) and (7) are useful In a sense that they significantly reduce the error[7], i.e. the difference between the simulated and the recorded data if the latter are non-stationary.

3. Model Fitting

The models (8) and (10) Which have been selected in the previous section for detection of structural of abnormalities in the time series (6), have the following parameters to be found from (6): the function F, the time delay T, the order of time delays m, the powers ml and m_2 of the difference $(1-B)^m$ and the seasonal difference $(1 - B')^m$ and the period S of the seasonal operator.

If the function F is linear, the simplest approach to model fitting is the Yule-Walker equations[4] which defines the autoregressive parameter α_i in Eq.(3) via the autocorrelations in (6). However, in many cases the assumption about linearity of the underlying dynamical system leads to poor model fitting, and therefore, it is more practical to assume from the beginning that F is a nonlinear (and still unknown) function. In such a posedness, probably, the best tool for model fitting is a feed-forward neural net which approximates the true extrapolation mapping by a function parametrized by the synaptic weights anti thresholds of the network' There is a rigorous proof[7] that any continuous function can be approximated by a feed-forward neural net with only one hidden layer, and that is why in this work a feed-forward neural net with one hidden layer was selected for the model (8) fitting. The model (8) is sought in the following form:

$$y(t) = \sigma \left\{ \sum_j W_{ij} \sigma \left[\sum_k w_{jk} \, y(t - k\tau) \right] \right\} \tag{11}$$

where W_{ij} and w_{jk} are constant synaptic weights, $\sigma(x) = \tanh x$ is a sigmoid function, and y(t) is a function which is supposed to approximate the stationarized time series (9) transformed from the original time series,

The model fitting procedure is based upon minimization of the error measure:,

$$E(W_{ij}, w_{jk}) = \frac{1}{2}\sum_{i,\mu}\left| y_i^\mu (t) - \sigma\left[\sum_j W_{ij}\sigma\left(\sum_k W_{jk}\, y(t-k\tau)\right)\right]\right\}^2 \tag{12}$$

where $y^\mu_i(t)$ are the values of the time series (9).

The error measure (12) represents the contribution of random components in the time series. There are two basic sources of such components. The first source is chaotic instability y of the underlying dynamical system; in principle, this component can be detected by applying the stabilization principle[5]. The second source is physical noise, imprecision of the measurements, or human factor (such as multi-choice decisions in economical or social systems). I-his component can be identified by applying a special type of dynamics (called terminal, or non-Lipschitz-dynamics)[6].

In this paper we will assume that E represents a variance of a mean zero, Gaussian noise.

Since there is an explicit analytical dependence between E and W_{ij}, w_{jk}, the first part of minimization can be preformed by applying back-propagation. However, further minimization should include more sophisticated versions of gradient decent since the dependence E (T, m, m_1, w_2, s) is too complex to be treated analytically.

4. Criterion of Structural Abnormalities

As noted in the introduction, there are two causes for abnormal behavior in the solution to Eq.(11): changes in external forces or initial conditions (these changes can be measured by Lyapunov stability and associated with operational abnormalities), and changes in the parameters W_{ij}, w_{jk}, i.e. changes in the structure of the function F in Eq.(8). (This changes are measured by structural stability and associated with structural abnormalities), and changes in the parameters W_{ij}, w_{jk}, i.e. changes in the structure of the function F in Eq. (8). (These changes are measured by structure stability and associated with structural abnormalities)

In this paper we introduce the following measure for structural abnormalities:

$$\zeta = \sum \left[(W_{ij} - \overset{o}{W}_{ij})^2 + (w_{jk} - \overset{o}{w}_{jk})^2 \right] \tag{13}$$

where $\overset{o}{W}_{ij}$ and $\overset{o}{w}_{jk}$, are the nominal, or "healthy" values of the parameters, and W_{ij}, w_{jk}, are their current values, Obviously, if

$$\zeta = 0, \ or \ \zeta < |\epsilon| \tag{14}$$

where ϵ is sufficiently small, then there is no structural abnormalities. The advantage of this criterion is in its simplicity: it can be periodically updated, and therefore, the structural "health" of the process can be easily monitored.

The only limitation of this criterion is that it does not specify a particular cause of an abnormal behavior. Obviously, this limitation can be removed by monitoring each parameter W_{ij}, w_{jk}, separately.

5. Application to Voyager Solar Pressure Data

The methodology presented in this paper has been applied to simulate the dynamics of Voyager solar pressure based upon the sensor data, The reference distribution of this data is shown in fig 1. The original data first required two simple differencing passes for stationarization (Fig.?). I-his data was fed then into the network at which point the feed-forward neural network learned the structural parameters of the system within two window iterations where a window consisted of approximately 100 data points. In this neural system, the simulated data was fitted into the stationarized data with a convergence? parameter value of 1 and a total of 20 neurons in the hidden layer. From fig. 3, we can see that the model simulations closely represent the actual Voyager data.

Since there is no apparent structural abnormalities in the original data, we therefore perform a procedure to test the sensitivit y for the parameter when faced with structural abnormalities. '1 his consists stretching in time the last third of the original data and running the altered data set through the neural system (Fig. 4). From Fig. 5, the structural change was immediately detected by a sharp change in the ζ parameter. Thus the system becomes highly sensitive for detecting structural abnormalities.

7. Concluding Remarks

A combined feed-forward neural-network and nonlinear dynamics approach to modeling stationary and non-stationar y time series was investigated. It was demonstrated that preliminary stationarization of the original data significantly improves the accuracy of simulations. A simple and reliable criterion of structural abnormalities was introduced and validated.

Acknowledgments

The research described in the paper was performed at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration in a collaborative program with the USCAR/Low Emissions Partnership (1 E P). Support also came from JPL's Multimission Ground Systems Office (MGSO). The authors wish to acknowledge useful discussions with Michael Hill, Ursula Schwutke and Robert Angelino.

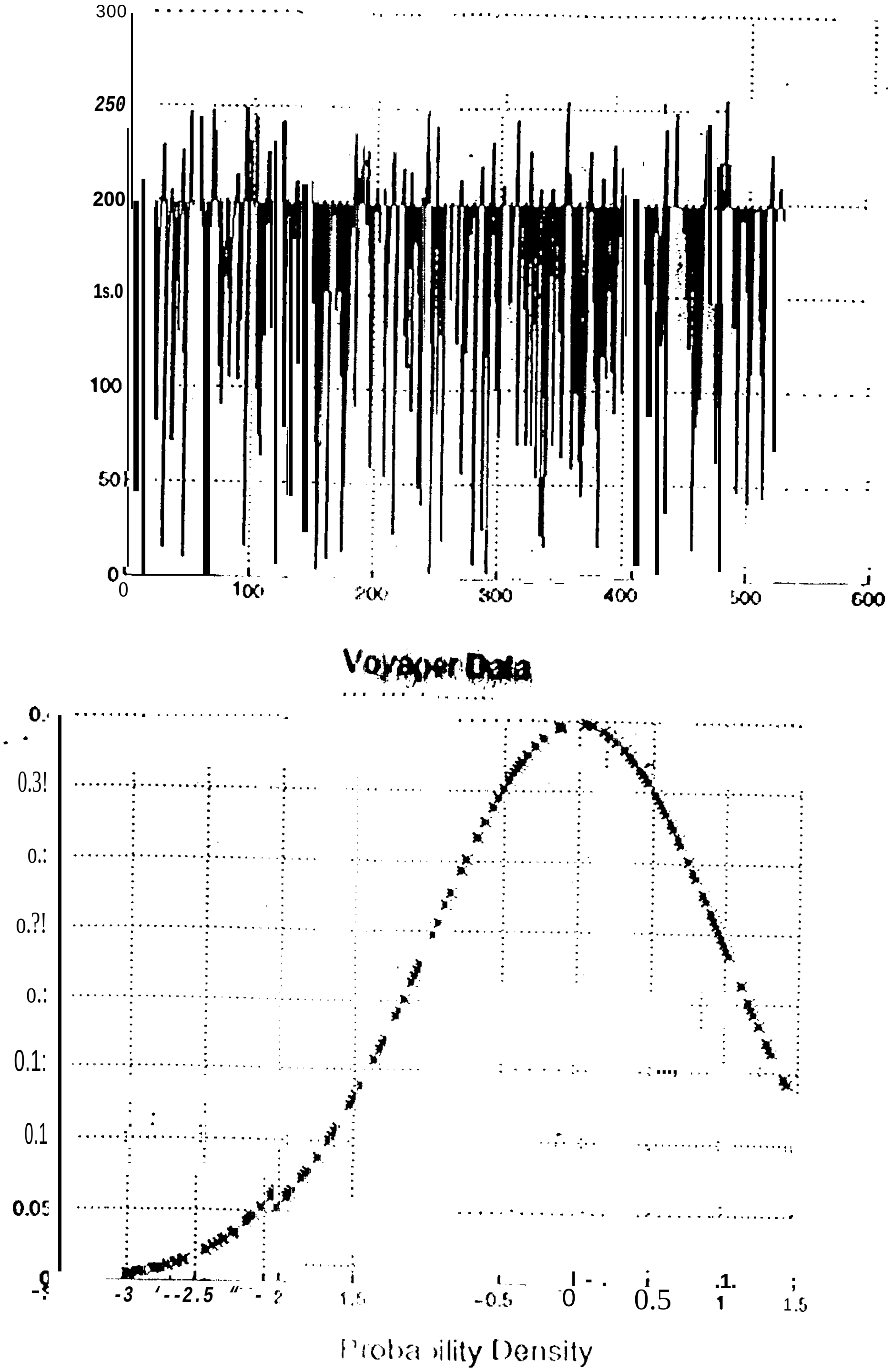


Figure 1

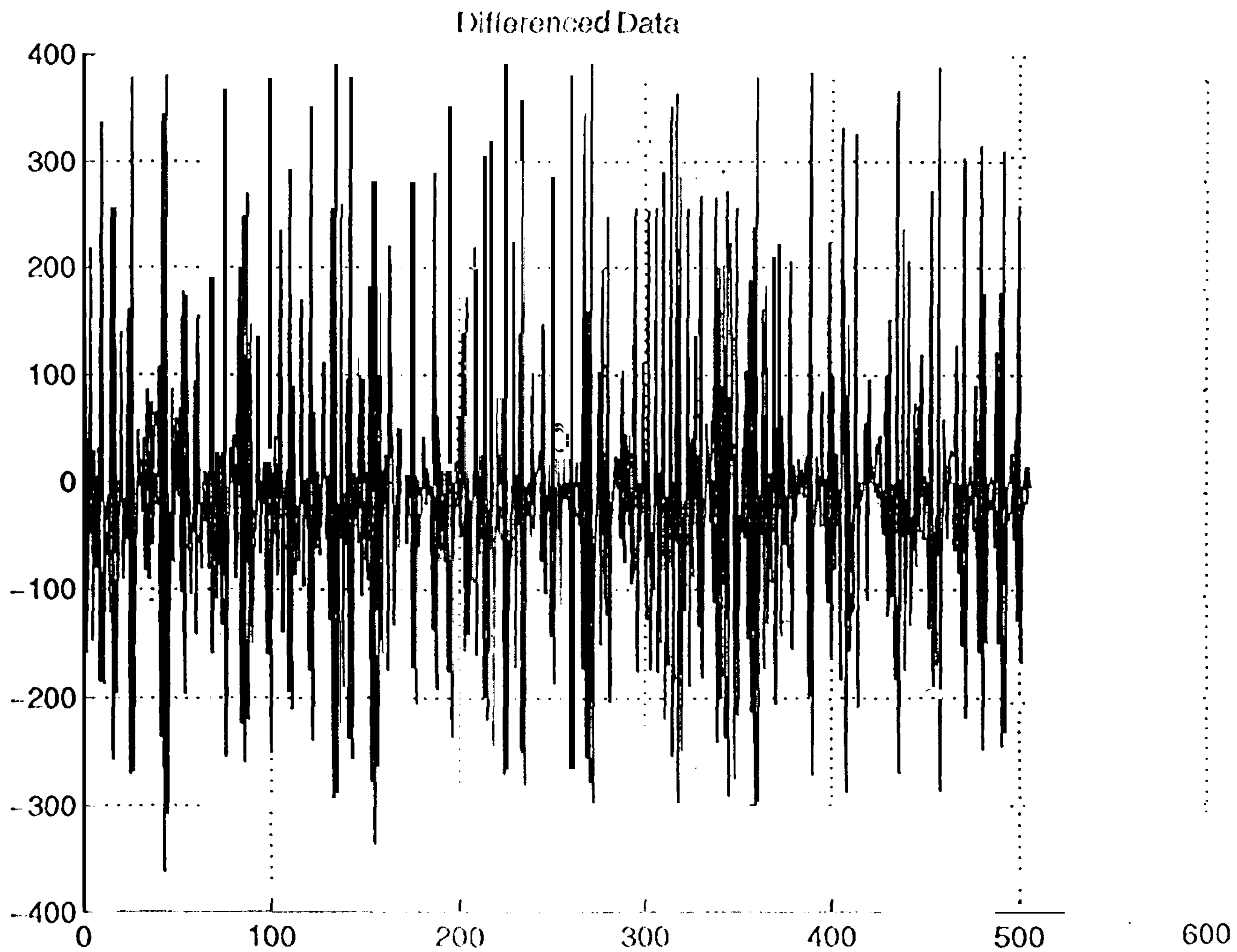
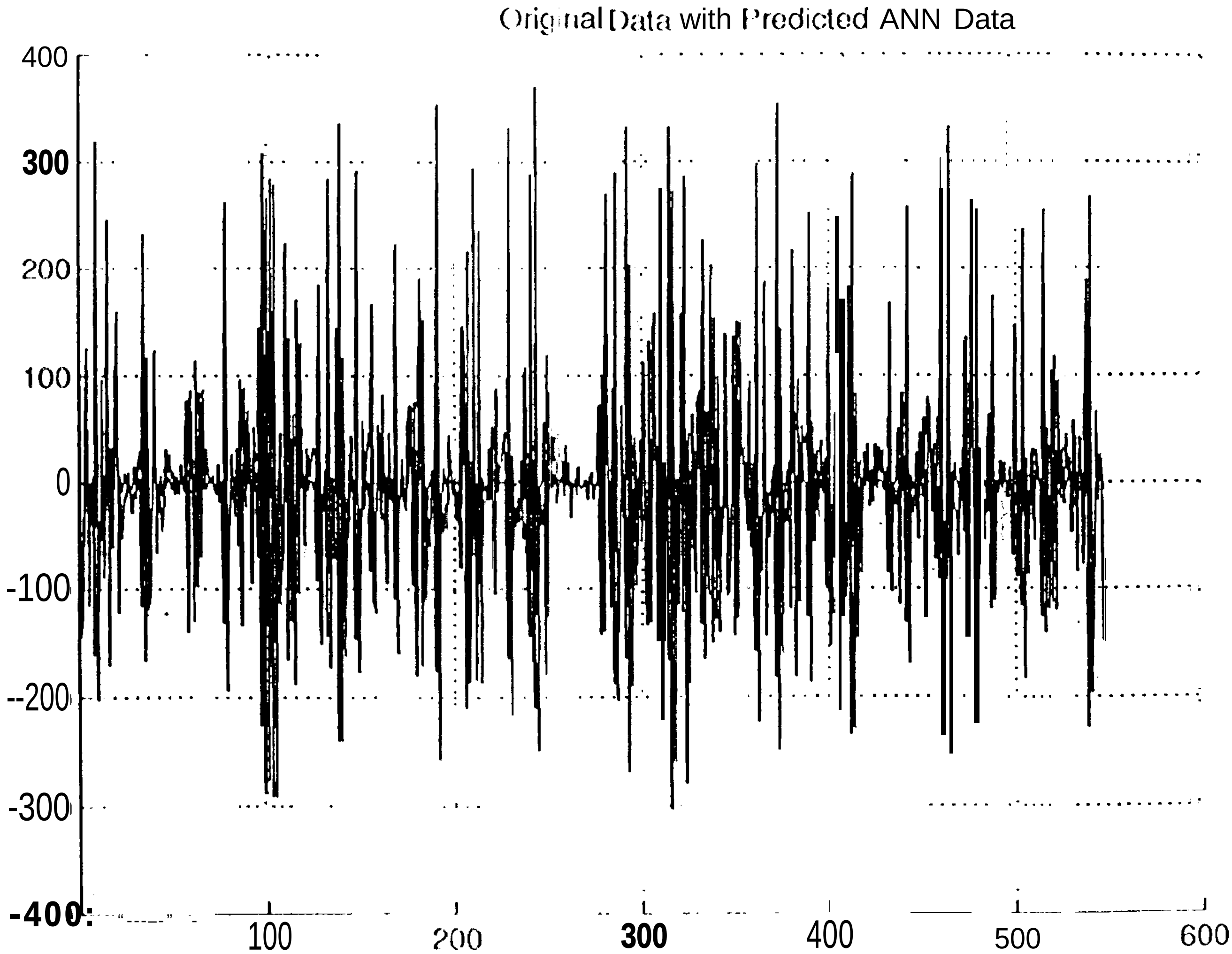
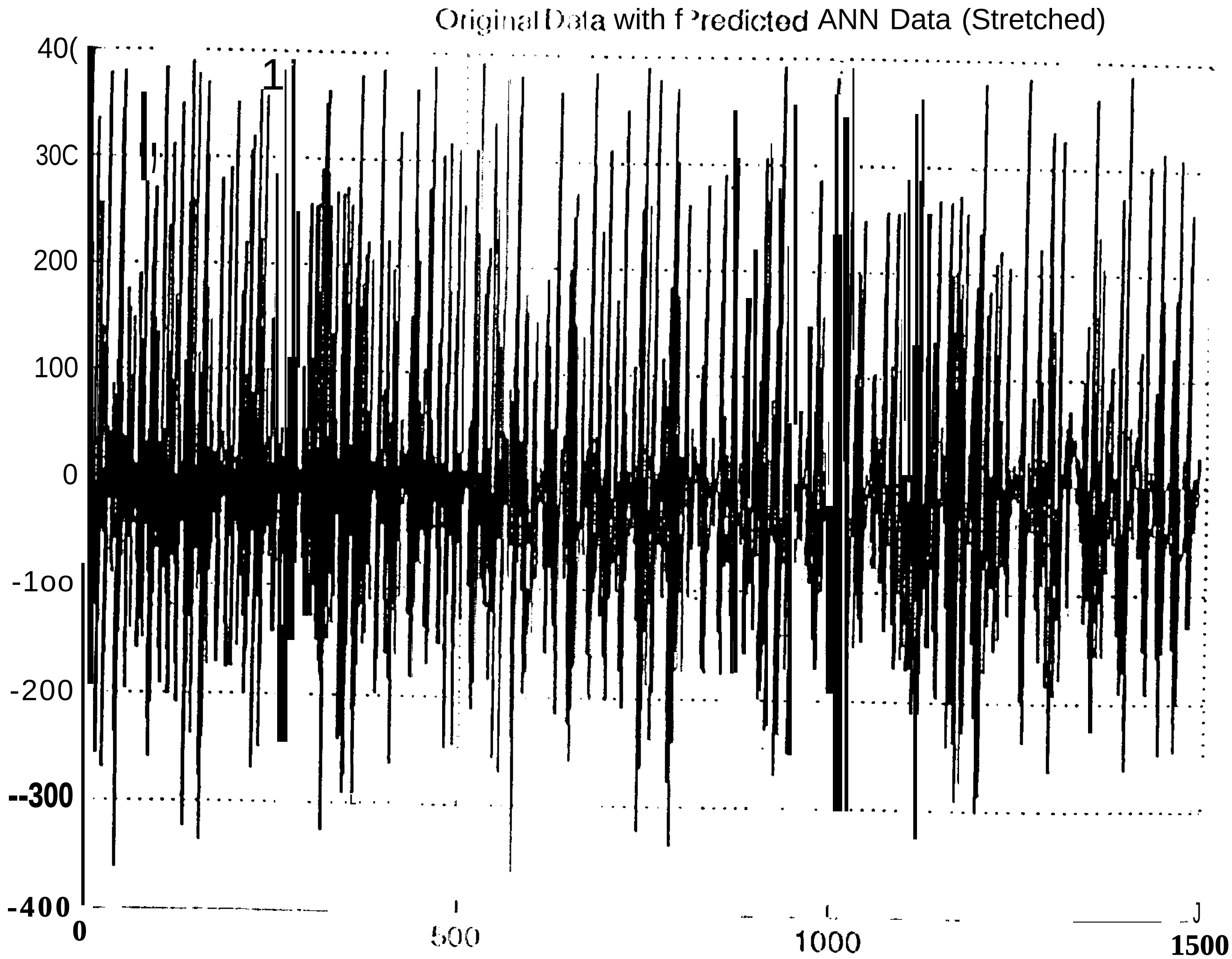

Differenced Data
400
300
200
100
0
-100
-200
-300
-400
0
100
200
300
400
500
600

Figure 2



Original Data with Predicted ANN Data
400
300
200
100
0
-100
-200
-300
-400
100
200
300
400
500
600

Figure 3



Original Data with Predicted ANN Data (Stretched)
400
300
200
100
0
-100
-200
-300
-400
0
500
1000
1500

Figure 4

References

1. Wheeler and A. Dhawan, C. Meyer, "SSME Sensor Modeling Using Radial Basis Function Neural Networks" AIM Journal of Spacecraft and Rockets.

2. C. Meyer, W. Maul "The Application of Neural Networks to the SSME Startup Transient", AIAA paper 91-2530

3. Takens, F., "Detection Strange Attractors on Turbulence" and "Dynamical Systems and "Turbulence", Lecture notes in Mathematics, Vol. 898 (Warwick 1980), Springer-Verlag

4. Box, E., "Time series analysis" Prentice Hall, NJ 1994.

5. Zak, M., Postinstability Models in Dynamics, Int. J. Of Theor Physics, Vol. 33, No. 77, 1994.

6. Zak, M., "Physical Models of Cognition" Int. 1. of Theor Physics, Vol. 33, No. 5, 1994.

7. Introduction to the theory of neural computations J. Hertz,...etc. Addision-Wesley Publishing Co. 1991.